Reconstruction, Analysis, and Segmentation of LA-ICP-MS Imaging Data using Python for the Identification of Sub-Organ Regions in Tissues

<table>
<thead>
<tr>
<th>Journal</th>
<th>Analyst</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>AN-ART-12-2019-002472.R1</td>
</tr>
<tr>
<td>Article Type</td>
<td>Paper</td>
</tr>
<tr>
<td>Date Submitted by the Author</td>
<td>12-Feb-2020</td>
</tr>
</tbody>
</table>
| Complete List of Authors | Castellanos-Garcia, Laura; University of Massachusetts Amherst, Chemistry
Elci, S.; Pamukkale University, Biomedical Engineering
Vachet, Richard; University of Massachusetts Amherst, Department of Chemistry |
Reconstruction, Analysis, and Segmentation of LA-ICP-MS Imaging Data using Python for the Identification of Sub-Organ Regions in Tissues

Laura J. Castellanos-García, S. Gokhan Elci and Richard W. Vachet

Laser ablation inductively coupled plasma mass spectrometry (LA-ICP-MS) imaging has been extensively used to determine the distributions of metals in biological tissues for a wide variety of applications. To be useful for identifying metal biodistributions, the acquired raw data needs to be reconstructed into a two-dimensional image. Several approaches have been developed for LA-ICP-MS image reconstruction, but less focus has been placed on software for more in-depth statistical processing of the imaging data. Yet, improved image processing can allow the biological ramifications of metal distributions in tissues to be better understood. In this work, we describe software written in Python that automatically reconstructs, analyzes, and segments images from LA-ICP-MS imaging data. Image segmentation is achieved using LA-ICP-MS signals from the biological metals Fe and Zn together with k-means clustering to automatically identify sub-organ regions in different tissues. Spatial awareness also can be incorporated into the images through a neighboring pixel evaluation that allows regions of interest to be identified that are at the limit of the LA-ICP-MS imaging resolution. The value of the described algorithms is demonstrated for LA-ICP-MS images of nanomaterial biodistributions. The developed image reconstruction and processing approach reveals that nanomaterials distribute in different sub-organ regions based on their chemical and physical properties, opening new possibilities for understanding the impact of such nanomaterials in vivo.

Introduction

Laser ablation inductively-coupled plasma mass spectrometry (LA-ICP-MS) imaging is increasingly used for imaging metal distributions in biological tissues. Among metal imaging techniques, LA-ICP-MS is perhaps the most sensitive technique for elemental imaging with detection limits in the sub µg/g level, while providing multiplexed metal analysis with spatial resolutions in the 10 to 200 µm range. Imaging methods such as scanning electron microscopy with energy-dispersive X-ray analysis (SEM-EDX) and particle-induced X-ray emission (PIXE) offer better spatial resolution; however, they require laborious sample preparation, and their sensitivity is lower than LA-ICP-MS. Synchrotron radiation X-ray fluorescence (SR-XRF) also offers better spatial resolution and has similar sensitivity to LA-ICP-MS, but it requires access to a synchrotron facility, making it much less broadly applicable. Given its combination of sensitivity, multi-metal detection capability, and accessibility, LA-ICP-MS has been used broadly in applications that include analysis of metals in neurogenerative diseases like Alzheimer’s, Parkinson’s, and Wilson’s disease, detection of anti-cancer metallodrugs, studies of metalloproteins, and analysis of nanomaterials in biological tissues. Several reviews have detailed the development and use of LA-ICP-MS imaging for analysing biological tissues.

Obtaining site-specific information about metal distributions in LA-ICP-MS imaging requires images to be reconstructed from the metal ion signals. In contrast to more widely used matrix assisted laser desorption/ionization (MALDI) MS imaging, relatively few approaches have been described for image reconstruction and statistical analysis. The program IMAGENA, which was developed by Osterholt et al., was one of the first software developed for visualizing LA-ICP-MS data. A similar program called HDIP was recently developed by Teledyne for image reconstruction of LA-ICP-MS images. While IMAGENA, HDIP, Iolite and its associated interfaces, such as monocle and bioLite, are versatile tools for reconstructing images, they are not open source and offer minimal tools for the statistical analysis of the resulting images. In contrast, software such as LA-iMageS, MapIT! and iQuant2 are open source programs that enable image reconstruction from LA-ICP-MS data via user-friendly graphical user interfaces, but they also have limited built-in statistical analysis tools. Other image reconstruction approaches, including those based on readily available software such as Microsoft Excel have also been described, although most have limited capability for the statistical analysis of the imaging data sets. For most existing software, image reconstruction is the principal aim. As the applicability of LA-ICP-MS grows, though, especially for applications such as nanomaterial-based drug delivery systems, more sophisticated image processing methods such as image segmentation for region of interest (ROI) analysis...
or other statistical analysis methods are needed to extract more information from reconstructed images.\textsuperscript{28} Image segmentation, in particular, is valuable for characterizing analyte signals in histologically relevant regions of a tissue, so that the underlying biochemistry and biology can be better understood. Deeper biological insight into MALDI-MS imaging data has been achieved with image segmentation algorithms,\textsuperscript{29,30} but to our knowledge analogous approaches have not been readily adopted in LA-ICP-MS imaging methods.

Here, we present an open source software written in Python for LA-ICP-MS imaging reconstruction that implements more advanced segmentation algorithms for classification of ROIs in LA-ICP-MS images. The use of Python for image analysis offers tremendous flexibility because of the numerous libraries accessible via the software for image visualization,\textsuperscript{31} matrix operations,\textsuperscript{32} statistical analysis,\textsuperscript{33} and even more complex tasks like multimodality imaging.\textsuperscript{34} Using code written in Python, we demonstrate that distinct sub-organ features can be automatically identified using different metal distributions to perform spatially aware segmentation analyses. As an application of these image segmentation approaches, we show that nanomaterials distribute in different sub-organ regions based on their chemical and physical properties. We believe the described software will benefit current and potential users of LA-ICP-MS imaging as it will make accessible more sophisticated image processing tools for more deeply understanding the biological ramifications of metal distributions in tissues.

**Materials and methods**

**Nanomaterial synthesis**

Different nanomaterials, including gold nanoparticles, nanozymes, nanocapsules and bismuth sulfide nanorods (Fig. S1\textsuperscript{+}), were provided by collaborators who synthesized them according to published protocols. Gold nanoparticles were synthesized according to the Brust-Schiffrin two phase method.\textsuperscript{35} Different ligand coatings, including ones with positively-charged (TMTA) and negatively-charged (TEGCOOH) functional groups were used.\textsuperscript{36} Nanozymes were synthesized using the method described by Rotello and co-workers.\textsuperscript{37,38} Nanocapsule synthesis was performed according to the protocol described by Rotello and co-workers,\textsuperscript{39–41} and the bismuth sulfide nanorods were synthesized according to the method developed by Gendelman and co-workers.\textsuperscript{42}

**Tissue Sections**

To obtain tissues for the imaging experiments involving gold nanoparticles, nanozymes and nanocapsules, female Balb/c mice (8-week-old) were injected with the nanomaterial of interest. After 24 h, the mouse tissues of interest were extracted, and flash frozen and then kept at -80 °C until used for MS imaging. All animal protocols involving the gold nanomaterials were approved by the UMass Institutional Animal Care and Use Committee (IACUC), which is certified to have met the ethical guidelines of the National Institutes of Health for handling laboratory animals for research. In all cases, tissues were sliced at 20 μm using a LEICA CM1850 at -20°C, and then deposited on regular glass slides. Hematoxylin and Eosin (H&E) staining on adjacent slices was performed using the Rapid Chrome frozen section staining kit (Thermo Fisher Scientific).

**LA-ICP-MS data acquisition**

LA-ICP-MS images were obtained on a CETAC LSX-213 G2 laser ablation system coupled with a Perkin Elmer NexION 300x ICP-MS. Unless otherwise specified, the following laser parameters were used: 50 μm spot size, 15 μm/s scan rate, 3.65 J laser energy, 10 Hz laser frequency, and a 10 s of shutter delay, which allows for full ablation of the tissue. The He carrier gas from laser ablation system was set to 0.6 L/min. The ICP-MS parameters were the following: 0.7 L/min nebulizer argon flow rate, 16.5 L/min plasma argon flow rate, 1.4 L/min auxiliary argon flow rate, -1650 V analog stage voltage, and 1000 V pulse stage voltage. These parameters were optimized for nanoparticle analysis in tissue sections, based on previous work.\textsuperscript{14,43,44} Different elements, including 197\textsuperscript{Au}, 209\textsuperscript{Bi}, 57\textsuperscript{Fe}, and 66\textsuperscript{Zn}, were detected with 50 ms dwell times. Image reconstruction and analysis was performed using a program written in Python that are further described below. Access to the scripts, examples and documentation can be found at: add github link.

**Normalization**

Normalization of the data allows the correction of tissue inhomogeneities during sample preparation and due to differences in mass ablation rates during laser ablation. Normalization of the tissues was performed using the Zn signals because we have empirically found that this element is relatively constant at the spatial resolutions and for the tissues used in this study (see Fig. S4). However, the user can perform normalization in the code using any desired element by simply specifying the element of interest as an input in the analysis workflow. Each of the studied metals (Au, Fe, and Bi) were divided by the Zn matrix in the Python script, on a pixel-by-pixel basis, using the NumPy library for matrix manipulations.\textsuperscript{32}

**Image segmentation and k-means clustering**

A k-means clustering protocol was performed in Python, using the scikit-learn machine learning library.\textsuperscript{33} To do this, the image of interest was vectorized, and the clustering was performed over the flattened image, as shown in Fig. S2. The number of
clusters was specified as a parameter in the program. If the number of clusters were unknown, the ‘elbow method’ was used to estimate the number of clusters into which the data should be divided, as shown in Fig. S3.\(^6\) After k-means clustering was performed, the centroids of the clusters were calculated, the data was reshaped, and the labelled image was generated. The labelled image corresponds to an arbitrary mathematical label that marks a specific area of the tissue as a part of a cluster.

## Results and discussion

### Image reconstruction

Generating an LA-ICP-MS image after data acquisition involves several data handling steps (Fig. 1). Because the metal signals are generated by scanning the laser in a line across the tissue, the continuously ablated and detected stream of material must be summed to generate an image pixel. To do this, the file that is generated for each ablation line, which contains all the metal intensity data, is subjected to a data reduction step (Step 2, Fig. 1). This step uses the laser spot size and scan rate to define the number of ion intensity measurements that are summed to create a pixel. For example, if the laser spot size is 50 µm and the laser scan rate is 15 µm/s, data acquired over a 3.3 s period is summed to create a single pixel. The resulting collection of pixels that contains ion intensity information is then separated into a set of different data matrices that correspond to the number of different metals measured (Step 3, Fig. 1). Separate images for each metal can then be reconstructed using plotting tools such as Matplotlib to generate a 2D image for each of the studied metals (Step 4, Fig. 1).\(^3\) This image generation approach has been written in a Python program that allows automatic image reconstruction with few user inputs. Our approach generates image matrices that can be subsequently analyzed by the many Python statistical libraries that exist, such as SciPy\(^4\) and scikit-learn enabling us to automatically identify different tissue regions in LA-ICP-MS images.

### Tissue boundary identification

The distribution of Zn signals from an imaging experiment can be used to delineate the edge of the tissue by differentiating the pixels that correspond to the tissue and those that correspond to the background. The procedure, illustrated in Fig. 2, requires measurements of background regions outside the tissue. In the first step of the procedure, the background Zn signal is calculated from any user-defined row or column in the imaging dataset. The row and column data are saved as two independent vectors, and the signal average and standard deviations in each case are calculated. The resulting average signals and standard deviations calculated are used to set the background value B. Each pixel in the entire image is then compared against the background and classified as tissue or background depending on whether its intensity is significantly different from the background signal, according to the equation in Fig. 2. From the classified image a background mask is generated in which tissue pixels are given a value of 1 and background pixels are 0. The approach used here is essentially a thresholding approach that is similar to that used by others,\(^4\) and in principle any element could be used by the program to perform background subtraction. Defining the tissue boundary is necessary for performing various statistical analyses on the images (see below).

### Image Optimization

In addition to helping define tissue boundaries, Zn signals can also help improve image quality in regions that are degraded by tissue inhomogeneities arising during sample preparation or from fluctuations in laser fluence or mass ablation rates. Zn is homogeneously distributed in many healthy tissues, such as liver, kidney and spleen, and in the tissues imaged in this work, Zn signals are empirically found to be homogeneously distributed as compared to other elements (Fig. 5A).\(^8\) This relative homogeneity allows Zn to be used for normalization. It should be noted, however, that any element, such as carbon\(^4\) or phosphorous,\(^15,16\) could be used for normalization in the program. Fig. 3 shows two examples of the advantage of using Zn signals for improving image quality. In Fig. 3a, wrinkling of the edges of a mouse spleen section causes the Au image to be poor throughout most of the tissue. When the Au intensity...
matrix is divided by the Zn matrix on a pixel-by-pixel basis, this normalization process improves the image by eliminating the zones with anomalously high overall Au signal that are caused by folding of the tissue edges. Similar improvements can be obtained for Fe images from a folded tissue (Fig. S5) and Au images from an experiment where the laser energy deviated during the experiment (Fig. 4b).

Image segmentation for automatic sub-organ differentiation

The distribution of Fe levels in a tissue depends on the blood flow to a specific sub-organ region and can be used to differentiate regions in various tissues.\textsuperscript{14,44,50} To distinguish sub-organ regions, image segmentation was performed using $k$-means clustering\textsuperscript{51} to partition areas of differential Fe composition. Fig. 4 shows H&E stained, optical and LA-ICP-MS images of liver and kidney sections from a mouse injected with TTMA and TEGCOOH Au nanoparticles (Fig. S1). While the areas of high blood flow (i.e. veins) are readily apparent from the H&E stained, optical, and Fe LA-ICP-MS images, image segmentation can be used to automatically identify these and other regions that are not as readily apparent. Using the Fe matrix as input, we performed $k$-means clustering with the number of clusters assigned as 3, based on the ‘elbow method’ (Fig. S3). When using $k$-means clustering, it is possible to effectively segment the image between background, low Fe (tissue), and high Fe (vein). Using image segmentation in this way allows one to determine the average signal of another metal in a given classified area. For example, the average Au signal can be determined in the three classified areas in liver and kidney sections from mice injected with TTMA (Fig. 4a) or TEGCOOH (Fig. 4b) Au nanoparticles. From the signal averages we can conclude that the Au nanoparticles accumulate differently in the liver and kidney. In the liver, we find more Au in the tissue than in the veins, and in the kidney, we find higher Au signal in the veins than in the rest of the tissue. Previous work by our group found that positively-charged nanoparticles like TTMA are readily cleared from circulation while negatively-charged nanoparticles like TEGCOOH circulate longer in the bloodstream\textsuperscript{44} which explains the differences in the nanoparticle concentrations in the veins of the two organs. This image segmentation approach allows this information to be automatically determined.
The image segmentation method was also used to distinguish sub-organ regions of the spleen. The spleen tissue has a marked difference between the red pulp and white pulp in that each region fulfills a different biological role in this vital organ. The spleen red and white pulp can be differentiated by their Fe concentrations, as the red pulp has higher blood flow than the white pulp. An example Fe LA-ICP-MS image from a spleen section is shown in Fig. 5a, showing areas of high and low Fe concentrations. To differentiate the sub-organ areas, we performed k-means clustering in the same manner as before. Using the elbow method, three clusters are again identified, but a k-means = 2 clustering and Zn-based background mask (B. Mask) determination (in brackets) produces a segmented image that accurately defines the red pulp, white pulp, and background. Neighboring pixel evaluation adds spatial awareness to the multimetal segmented image by redefining each labeled image pixel \( P_{n,m} \) to a weighted image pixel \( WP_{n,m} \) using the indicated equation. The result is a weighted image that clearly defines the boundary between the red and white pulp, allowing differentiation of the marginal zone that separates the two regions. White scale bars in Fe image correspond to 500 µm.

The image segmentation method was also used to distinguish sub-organ regions of the spleen. The spleen tissue has a marked difference between the red pulp and white pulp in that each region fulfills a different biological role in this vital organ. The spleen red and white pulp can be differentiated by their Fe concentrations, as the red pulp has higher blood flow than the white pulp. An example Fe LA-ICP-MS image from a spleen section is shown in Fig. 5a, showing areas of high and low Fe concentrations. To differentiate the sub-organ areas, we performed k-means clustering in the same manner as before. Using the elbow method, three clusters are again identified, but using a k-means clustering of three does not allow an effective differentiation between the white pulp and the background primarily because of the spread of Fe signals in the red pulp exceeds the difference between the white pulp and the background Fe signals (Fig. 5a, k-means = 3 image). To solve this issue and effectively differentiate the background and white pulp regions, we used a multi-metal segmentation strategy that is illustrated within the brackets of Fig. 5a. For this strategy, we choose a k-means cluster value of 2 for the Fe image to differentiate two clusters, one exclusively for the red pulp and another for the white pulp and background. The white pulp and background are then differentiated in the k-means = 2 clustered image using the Zn signal and the background mask procedure illustrated in Fig. 2. By conjugating the k-means = 2 clustered image and the background mask (Fig. 5a), we can generate a multi-metal segmented image with three distinctive areas: background, red pulp, and white pulp.

In addition to the distinct red and white pulp regions of the spleen, there is a boundary region known as the marginal zone where the first steps of an immune response occur in this organ. Image segmentation alone makes it difficult to effectively differentiate the marginal zone because it does not have a distinct metal composition. Because the marginal zone surrounds each white pulp region and is approximately 50 µm in size, this region can be distinguished if spatial awareness is added to the segmented image. The k-means approach, however, is performed on a vectorized dataset and thus does not have spatial awareness.

Spatial awareness can be added by considering the neighboring pixels around any particular pixel in the image by arbitrarily assigning values of 0, 1, and 2 to the background, red pulp, and white pulp pixels, respectively, that were identified via the multi-metal segmentation approach (see label image in Fig. 5b).
To further classify distinct areas in the spleen, including the marginal zone, each pixel value or label ($P_{n,m}$) can be redefined as a weighted pixel ($WP_{n,m}$) that is equal to the weighted average of its eight immediately neighboring pixels (see equation in Fig. 5b). In effect, we apply image filtering with a linear filter to classify the boundary regions of the labeled image. After redefining the value of each pixel, we can then generate a weighted image that effectively distinguishes the marginal zone (in yellow) from the red and white pulp (see weighted image in Fig. 5b). Newly weighted values of 1.2-1.3 correspond to the marginal zone, while lower and higher values correspond to the red and white pulp, respectively. This approach for distinguishing the red pulp, white pulp, and marginal zone can be validated by comparing an H&E stain of the spleen tissue with the resulting multi-metal segmented image (Fig. S7 and S8). The red and white pulp stain differently, and the overlaid images show excellent agreement despite the images coming from adjacent slices. It should also be noted that the neighboring pixel approach illustrated in Fig. 5b can also correct wrongly classified regions in segmented images of other tissues (see Fig. S9).

### Gold nanoparticle and bismuth sulfide nanorod distributions in spleen tissues

The value of distinguishing the three different regions in the spleen using multi-metal segmentation and the neighboring pixel evaluation can be illustrated by considering LA-ICP-MS images from adjacent slices. It should also be noted that the neighboring-pixel evaluation procedure introduces spatial awareness to the segmentation process that can correct for misclassified pixels and can classify boundary regions that are at the limit of the measurement resolution (e.g. marginal zone in the spleen).

### Conclusions

We have developed software written in Python that can automatically reconstruct and segment images from LA-ICP-MS imaging data. This new software identifies sub-organ regions of interest with minimal user input and can find regions that might be missed by manual analysis. The image reconstruction program takes advantage of the capability of existing open source scientific libraries such as NumPy, Matplotlib, and Scikit-learn for various numerical and statistical analyses. Our image reconstruction and analysis method represents the first open-source software, to our knowledge, that can perform sophisticated manipulations automatically and directly on LA-ICP-MS imaging data. Using this software, we demonstrate that segmentation of LA-ICP-MS images can be performed using a combination of Fe and Zn images, k-means clustering analysis, and neighboring-pixel evaluation to automatically classify sub-organ regions in kidney, liver, and spleen tissues. The neighboring-pixel evaluation procedure introduces spatial awareness to the segmentation process that can correct for misclassified pixels and can classify boundary regions that are at the limit of the measurement resolution (e.g. marginal zone in the spleen). Using tissues from mice injected with different nanomaterials as examples, classification of different sub-organ...
regions reveals the value of our described approach. For example, we find that Bi sulfide nanorods accumulate more extensively than Au nanoparticles in the marginal zone as compared to other regions of the spleen. We believe that the described data reconstruction and image segmentation strategy that we have developed in Python will be beneficial to LA-ICP-MS imaging experts and non-experts alike. Moreover, the use of Python allows a wide array of other statistical methods to be applied to the data taken during an LA-ICP-MS imaging experiment. We envision future developments of the code that include the incorporation of standards into the workflow to allow for quantitative imaging, Pearson’s coefficient calculations for determining co-registration extents of different data channels, and outlier detection in ROIs.

Conflicts of interest

In accordance with our policy on Conflicts of interest please ensure that a conflicts of interest statement is included in your manuscript here. Please note that this statement is required for all submitted manuscripts. If no conflicts exist, please state that “There are no conflicts to declare”.

Acknowledgements

This research was supported by the National Science Foundation, Grant CHE-1808199. We thank Professor Vincent Rotello and researchers (Xianzhi Zhang, Roberto Cao-Milan, Yuanchang Liu, Yi-Wei Lee) in his lab at the University of Massachusetts Amherst for the nanoparticle samples. We also thank Prof. Howard Gendelman and his group from the University of Nebraska for the Bismuth nanorods and mouse tissue samples. Microscopy experiments were performed in the Nikon Center of Excellence, which is part of the Institute for Applied Life Sciences at University of Massachusetts Amherst.

Notes and references

‡ Footnotes relating to the main text should appear here. These might include comments relevant to but not central to the matter under discussion, limited experimental and spectral data, and crystallographic data.


29 T. Alexandrov, M. Becker, S.-O. Deininger, G. Ernst, L.


389x83mm (96 x 96 DPI)